PARADOXES OF INFINITE AGGREGATION

Frank Hong
The University of Hong Kong

Jeffrey Sanford Russell
University of Southern California

July 2024

Abstract

There are infinitely many ways the world might be, and there may well be infinitely many people in it. These facts raise moral paradoxes. We explore a conflict between two highly attractive principles: a *Pareto* principle that says that what is better for everyone is better overall, and a *statewise dominance* principle that says that what is sure to turn out better is better on balance. We refine and generalize this paradox, showing that the problem is faced by many theories of interpersonal aggregation besides utilitarianism, and by many decision theories besides expected value theory. Considering the range of consistent responses, we find all of them to be quite radical.

1 The Good and the Infinite

There are many ways the world might be: indeed, infinitely many. Evaluating how good an uncertain prospect is depends on how the value spread across infinitely many states contributes to the prospect's overall value.

There are many people in the world: indeed, there may well be infinitely many.[1] Evaluating how good a world is overall for all of the people in it

[1] Physicists take very seriously the hypothesis that the universe has infinite spatial volume (see Askell, 2018, chapter 1). Supposing it does, on natural assumptions it is almost sure that that there are infinitely many people.

depends on how the value spread across infinitely many individual people contributes to the overall value for that population.

Much has been written on each of these two questions, but explorations of how the two questions interact have only just begun. (For overview, see Easwaran et al., 2021, sec. 7.2; Arntzenius, 2014, and references therein.)

Let's start with a story.[2]

> *The Social St. Petersburg Paradox.* Hilbert has opened a grand new hotel in St. Petersburg. As with all of Hilbert's hotels, this hotel has one room for each natural number. Each room is occupied by one guest. These guests are epicures who are capable of remarkable amounts of happiness; moreover, happiness is good for them precisely in proportion to its amount. Hilbert is considering two options for how to allocate the impressive resources at his disposal to make his guests happy.
>
> 1. *Equal and Risky.* Every guest gets to play a local favorite casino game: the *St. Petersburg game*. A fair coin is flipped until it comes up heads. If the first heads is on the nth flip, then everyone gets 2^n hedons (units of happiness).
> 2. *Unequal and Safe.* The person in the first room gets 2 hedons for sure, the person in the second room gets 4 hedons for sure, and so on, with the nth person getting 2^n hedons for sure.

These two options are summarized in Table 1. Each column H_k represents the event in which the first heads is on the kth flip. The nth row represents the amount of happiness the person in room n receives in each state H_k.

[2]This story builds in strong and unrealistic simplifying assumptions—for instance about numerical "units of happiness"—many of which will be relaxed in what follows. The argument in this section is closely related to an argument previously given by Wilkinson, 2022; see also Goodsell, 2021. We discuss this further at the end of this section. It is also structurally similar to Cain (1995)'s "sphere of suffering" argument, which shows a conflict between principles about aggregating across people and aggregating across times. The main way the Social St. Petersburg argument goes differently is because we have to take probabilities into account, which do not have a direct analogue for times. The more powerful argument in Section 2 requires further departures from Cain's setup.

Equal and Risky	H_1	H_2	H_3	$\cdots$
Person in room 1	2	4	8	$\cdots$
″ 2	2	4	8	$\cdots$
″ 3	2	4	8	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Unequal and Safe	H_1	H_2	H_3	$\cdots$
Person in room 1	2	2	2	$\cdots$
″ 2	4	4	4	$\cdots$
″ 3	8	8	8	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Table 1: The Social St. Petersburg Paradox

Which option is better for Hilbert's guests, on balance? There are strong arguments either way.

Argument for Equal and Risky. Compare the two tables row by row. The person in room n faces a choice between playing a St. Petersburg game for happiness, or else getting a certain finite quantity of happiness, 2^n hedons, for sure. There are strong arguments that the St. Petersburg game is better. One argument is based on this principle:[3]

Individual Expectations. For any prospects X and Y and individual i, if the expected amount of goodness of X for i is greater than the expected amount of goodness of Y for i, then X is better for i than Y.

For Equal and Risky, the expected amount of goodness diverges to positive infinity:

$$\frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 4 + \frac{1}{8} \cdot 8 + \cdots = 1 + 1 + 1 + \cdots$$

For Unequal and Safe, the expectation is finite. This looks like a good reason to say that Equal and Risky is better than Unequal and Safe for each

[3]Broome (1991, pg. 142) calls this *Bernoulli's Hypothesis*.

person. (The argument just given is not airtight—we will improve on it later.)

If Equal and Risky is strictly better *for each person*, this seems like a powerful reason for thinking that it is strictly better overall—appealing to the following principle:

Ex Ante Pareto. (Weak) If X is at least as good as Y for every individual, then X is at least as good as Y overall.

(Strict) If X is strictly better than Y for every individual, then X is strictly better than Y overall.

Making everyone better off, individually, is a way to make things better for everyone, collectively.

Argument for Unequal and Safe. Compare the two tables column by column. In each state H_k, Equal and Risky results in a certain finite quantity of happiness, 2^k hedons, for every individual. Unequal and Safe, meanwhile, results in an exponential distribution of happiness—and while this distribution is highly unequal, it is extremely good for many people. Consider $k = 3$. Then Equal and Risky gives each person 8 hedons. Unequal and Safe makes two people worse off than this—each by less than 8 hedons. Meanwhile *infinitely* many people are better off—each by *at least* 8 hedons. So it seems very plausible that the unequal distribution is better overall. Similar considerations apply to every state H_k: Unequal and Safe makes only $k-1$ people worse off, each by less than 2^k hedons, and *infinitely* many people better off, each by at *least* 2^k hedons. This reasoning suggests that however the coin flips go, Unequal and Safe has better results overall than Equal and Risky.

One general principle that would underwrite this judgment is *Additivity* (see Vallentyne & Kagan, 1997; Lauwers & Vallentyne, 2004; Basu & Mitra, 2007; Wilkinson, 2022). This is a natural infinite generalization of the *utilitarian* principle, which says that the good for a population is given by the sum of the goods for each individual. We can't apply this principle directly, since the sum of goods in each distribution is infinite. The idea of the generalization is that, if the total amount by which people are made better off by switching from outcome x to outcome y is greater than the total improvement from switching from y to x, then y is better than x overall. Let the *step up* from x to y be the sum of the difference in value between x and y for each individual who is better off in y than they are in

x. If the step up from x to y is greater than the step up from y to x, then y is better than x.[4]

In the case at hand, in each state H_k, the step up from Unequal and Safe's outcome to Equal and Risky's outcome is finite. (In fact, it's less than $k \cdot 2^k$.) But the step in the opposite direction is infinite. So Additivity tells us that Unequal and Safe turns out better than Equal and Risky. Since the same reasoning goes for every state, Unequal and Safe is sure to turn out better than Equal and Risky.

This seems like a powerful reason to say that Unequal and Safe is a better prospect than Equal and Risky—appealing to the following principle.

Statewise Dominance. (Weak) If the outcome that results from a prospect X is at least as good as the outcome that results from a prospect Y in every possible state, then X is at least as good as Y.

(Strict) If the outcome that results from a prospect X is strictly better than the outcome that results from a prospect Y in every possible state, then X is strictly better than Y.

Sometimes comparing two uncertain prospects is hard: we may not be sure which of them will turn out better, and we have to weigh up the different possible outcomes. But sometimes we *are* sure which prospect will turn out better—and in this case, comparing them should be easy.

To sum up, "personwise" reasoning based on Ex Ante Pareto tells us that Equal and Risky is better than Unequal and Safe, while "statewise" reasoning based on Statewise Dominance tells us that Equal and Risky is worse than Unequal and Safe. The two lines of reasoning cannot both be correct.

The Social St. Petersburg Paradox is relatively easy to state, but it relies on strong, controversial assumptions. We have assumed that every individual's good can be measured on a common numerical scale ("hedons"), and we have made strong assumptions both about how these numerical quantities of goodness are to be aggregated across states (Individual Expectations) and across individuals (Additivity). Each of these views is mainstream enough that the clash between them is noteworthy. Indeed,

[4] This should be understood as including the case where the step up from x to y is infinite and the opposite step is finite. But we do *not* include the case where both steps are infinite.

the conflict between Individual Expectations, Additivity, Ex Ante Pareto, and Statewise Dominance has previously been noted by Wilkinson (2022).[5]

But there are many rival views about weighing risks, and there are many rival views about interpersonal aggregation. For example, some theorists recommend risk-aversion rather than risk-neutrality; others recommend discounting small enough probabilities; and besides standard utilitarians there are also prioritarians and egalitarians of various stripes. So it appears that there are many avenues for escape. In Section 2 though, we will find that the core problem is in fact more general. Much weaker assumptions about risk, interpersonal aggregation, and the structure of individual good suffice to raise a paradox; so strikingly diverse views must reckon with it.[6] In Section 3 we will canvas the directions a positive theory might still go. But the conflict between "personwise" and "statewise" reasoning runs deep.

2 Generalizing the Paradox

The Social St. Petersburg Paradox in Section 1 turns on a *numerical measure* of goodness, an *expectational* theory of risk, and an *additive* theory of interpersonal aggregation. We can substantially weaken all of these assumptions. Toward that, let's first consider another story.

Two Teams.[7] The hotel's epicurean guests (who were capable

[5]We have discussed the Social St. Petersburg paradox, rather than Wilkinson's (2022) "Egregious Energy" argument, mainly because it is simpler. But we should note that Wilkinson's version does have an important advantage. The Social St. Petersburg Paradox relies on the assumption that each person's goodness is unbounded: the "epicures" of the story can enjoy arbitrarily large quantities of good. For these people, any good can be compensated for by an arbitrarily small chance of something even better. Wilkinson's argument dispenses with that assumption, and instead uses the assumption that individual goods can be arbitrarily *small*. The general argument we present in Section 2 makes neither of these assumptions.

[6]Wilkinson (2022, footnote 1, p. 1) recognizes the possibility of such generalizations, though he does not explore this: "This definition excludes narrow person-affecting views, as well as any theory under which value does not admit an additively separable representation (e.g., egalitarianism, maximin, averagism). This exclusion is not because such views escape the infinitarian worries described below—typically, they dont—but simply for brevity."

[7]We have learned that Kowalczyk (n.d.) independently gives a very similar argument to the following. Goodsell (2021) gives a related argument in a different context.

of intense happiness) have all left, and a new group of people checks in to Hilbert's hotel. These people are incapable of enjoying the benefits of the extremely high stakes casino games that St. Petersburg traditionally offers, but there is something here for them too. First, Hilbert's guests are split into two infinite teams: Red Team and Blue Team. Each team is further divided into infinitely many larger and larger cohorts: for each team, the nth cohort includes 2^{n^2} people.

There are two new options (Table 2).

1. *One Team Wins.* We start by flipping a coin to decide which team gets to play: heads, Red; tails, Blue. Call this the *Team-Choosing Flip.* Then, as usual, we flip a coin until it first lands heads. Call these coin flips the *St. Petersburg Flips.* If the first heads is on the nth flip, then everyone in cohort n of the winning team gets a cake.
2. *Both Teams Can Win.* The same coin flips are used, but in a different way. The Team-Choosing Flip is ignored. If the *first* St. Petersburg Flip comes up heads, then no one gets any cake. Otherwise, if the first heads is on the $(n + 1)$th St. Petersburg Flip, then everyone in cohort n of *both* teams gets a cake.

Both games give each person in the nth cohort of either team probability $1/2^{n+1}$ of getting a cake, and otherwise nothing. So both options look equally good to everyone. Ex Ante Pareto says both games are equally good.

But also, no matter what happens, more people get cake in One Team Wins than in Both Teams Can Win. The Team Choosing Flip makes no difference to how many people win a cake. If the first St. Petersburg Flip comes up heads, then One Team Can Win gives two people cakes, while Both Teams Can Win gives nothing. If the $(n + 1)$th St. Petersburg Flip is the first heads, then One Team Wins delivers $2^{(n+1)^2}$ cakes, while Both Teams Can Win delivers 2^{n^2+1} cakes, which is less. Statewise Dominance says that One Team Wins is strictly better than Both Teams Can Win.

Now let's generalize this paradox.

One Team Wins	HH_1	TH_1	HH_2	TH_2	$\cdots$	HH_n	TH_n	$\cdots$
2 people	1							
2 people		1						
16 people			1					
16 people				1				
$\vdots$					$\ddots$			
2^{n^2} people						1		
2^{n^2} people							1	
$\vdots$								$\ddots$

Both Teams Can Win	HH_1	TH_1	HH_2	TH_2	HH_3	TH_3	$\cdots$	HH_{n+1}	TH_{n+1}	$\cdots$
2 people			1	1						
2 people			1	1						
16 people					1	1				
16 people					1	1				
$\vdots$							$\ddots$			
2^{n^2} people								1	1	
2^{n^2} people								1	1	
$\vdots$										$\ddots$

Table 2: Two Teams

First, we drop the assumption that how good outcomes are for each individual can be represented on a common numerical scale. Instead of numbers, we will abstractly consider allotments of **individual goods**. Think of a "good" as specifying everything about a person's life that contributes one way or another to their well-being. We do not assume that the goods for one individual are the *same* for another, nor that different individuals' goods are directly comparable. But we do need to make *some* interpersonal trade-offs: the paradox only arises if we suppose that losses for some people can be *compensated* by gains for other people. Let's make this idea precise. (The definitions that follow are made more precise in Appendix A.)

It will be convenient to fix a baseline "zero" level of good for each individual i, which we'll denote 0_i. The zero level does not need to have any special meaning—for example, it doesn't have to be a "life only barely worth living" filled with "muzak and potatoes" (as Parfit, 1986 memorably puts it). Think of each person's "baseline" as representing a pretty good life, including ordinary pleasures and fulfilment, without inordinate suffering. (Again, we don't need to identify one individual's baseline good with any other's.)

A **social outcome** determines some good for each individual: we'll write $x(i)$ for the good that social outcome x assigns to individual i. We call a social outcome x an **allocation (only) for** a set of individuals I iff $x(i) = 0_i$ for every individual i *not* in I.

Unlike the Social St. Petersburg Paradox, *Two Teams* did not rely on as strong a principle as Additivity. But we did still suppose that a cake is equally good for everyone, and that all that matters in the end is how many cakes get distributed. These axiological assumptions are still strong. We can replace them with a much weaker principle about interpersonal aggregation.

Interpersonal Compensation. For any allocation x for a finite set of individuals I, there is an allocation y for some finite set of individuals J disjoint from I such that y is better overall than x.

The idea is that whatever the I-individuals might sacrifice, by getting their baseline good rather than something else that might be better, can be compensated for somehow by improving the lot of some other group of individuals J.

The generalized paradox also relies on a principle about intrapersonal trade-offs involving risk. But it is nowhere near as strong as Individual Expectations, which was used in Section 1.

An *individual prospect* is a function from states to individual goods, while a *social prospect* is a function from states to social outcomes. For an event E and a good x for an individual i, let $x|_E$ be the individual prospect that has outcome x in event E, and the baseline outcome 0_i otherwise.

Stochastic Compensation. For any good x for an individual i, and any events E and F such that $P(E) \leq P(F)$, there is some good y for i such that $y|_F$ is at least as good for i as $x|_E$.

Stochastic Compensation is so-called because it is analogous to the Compensation principle between individuals: we can compensate for losing something in some states as long as we gain enough in other sufficiently probable states.[8] It is a very weak version of the widely accepted principle of *stochastic dominance* (see Tarsney, 2020, and references therein).

Now we can state our main result.

Theorem 1. *(Weak) Ex Ante Pareto, (Strict) Statewise Dominance, Interpersonal Compensation, and Stochastic Compensation are jointly inconsistent.*

The proof is a direct generalization of the reasoning in *Two Teams*; it can be found in Appendix A.

Now let's discuss each of the two principles we introduced in this section in more detail.

Interpersonal Compensation can be derived from the utilitarian picture where everyone's good is on a common numerical scale, and general good is given by adding things up. But it does not rely on that picture: Compensation is much weaker than Additivity.

Consider a famous case from Scanlon (1998, p. 235): an electrician suffers a horrible electric shock in order to keep a broadcast working to provide entertainment to vast multitudes. Additivity implies—counterintuitively—that if the multitudes are vast enough, then this is better than if the electrician escaped unharmed and the broadcast failed. But Interpersonal Compensation does not imply this. First off, recall that the "baseline" zero

[8]One may wish to restrict Stochastic Compensation to the case where $P(E) < P(F)$, to avoid the case where E is a proper subset of F but $P(F \setminus E) = 0$. This weakening suffices for all our applications, with small modifications to the proofs.

level is supposed to be pretty good: we can stipulate that it does not involve painful electrocution. But it may seem similarly counterintuitive to think that the electrician ought to forego some great benefit for the sake of the multitudes: say, giving up the chance to accomplish a lifelong ambition of hiking the Pacific Crest Trail. Interpersonal Compensation does not have this implication either. Unlike Additivity, Interpersonal Compensation is compatible with the view that the electrician's sacrifice would only be good if it spared many *more* people from much *worse* sacrifices. While it is counterintuitive that it is better for the electrician to gravely suffer than for many others to lose some minor entertainment, it is not so counterintuitive to think that it would be better for the electrician to lose a great benefit than for many others to lose even greater benefits.[9]

Similarly, unlike Additivity, Interpersonal Compensation is consistent with the view that equality matters intrinsically. For all the principle says, perhaps losses can only be compensated by benefits that are distributed at least as equally as what was lost.[10]

Furthermore, Additivity requires that individual good has a great deal of structure. Differences in individual good must be the sort of thing that can be *added*. This requires that each individual's good has a *cardinal* scale. But Additivity also requires us to add a difference in *one* person's individual good to a difference in someone else's good. This requires that all individuals' goods are on *co-cardinal* scales. (This means that we can fix numerical representations for each individual's good which are unique up to positive linear transformations which have the same constant scaling factor for every individual.) In contrast, Interpersonal Compensation assumes much less structure. It does not require that individual good can be represented by a real number, it does not require any arithmetical operations on individual good, and it requires very little in the way of comparability between what is good for one individual and what is good for another. Interpersonal Compensation does still require a little bit in the way of interpersonal comparisons: for example, it rules out the *Pure Paretian* view, according to which the *only* case where one social outcome

[9]That said, Horton (2017, sec. II. A.) argues that "limited anti-aggregationism" is an unstable position; his argument is related to our concerns here about the social distribution of risks.

[10]Note, though, that some implementations of the egalitarian idea are inconsistent with the Pareto principle. Note also that it is not entirely clear in general how to understand distributive equality for infinite populations.

x is at least as good overall as another y is when x is at least as good as y for every individual. But it doesn't require very much.

Now let's turn to what *Stochastic Compensation* says about weighing risks. In Section 1 we gave a simple argument using the Individual Expectations principle, in the course of arguing that a St. Petersburg lottery was better for each individual than any fixed finite good. As we noted there, this left the door open to many alternative ways of weighing risks that might avoid this counterintuitive "fanatical" consequence—and perhaps thereby escape the paradox. But Theorem 1 closes this door. Stochastic Compensation follows from the simpler principle that getting a good thing x with chance p is at least as good as getting x with chance less than p. So, for example, unlike Individual Expectations, Stochastic Compensation is satisfied by familiar theories of risk aversion (such as Buchak, 2013). It is also satisfied by many decision theories that discount small probabilities (as is recommended by Smith, 2014; Monton, 2019). This will include any theory where discounting is *monotone*—so a larger probability of a good outcome is never counted for *strictly less* than a smaller probability of that same outcome. Even a very small chance at cake will be *at least as good* as an even smaller chance at the same cake. Some discounting views will say that both prospects are *equally* good, if both probabilities are negligibly small—but that does not violate the principle.

Not all discounters neglect very improbable states entirely: they may still take them into consideration at least as tie-breakers.[11] Beckstead and Thomas (2021) propose an elegant formulation of such a theory, in which

[11]Monton (2019, p. 20) recommends this in reply to worries about dominance raised by Isaacs (2016). He also notes that some more orthodox decision theorists take recourse to a similar two-stage theory, since they have their own difficulties with dominance when probability zero events are involved.

It is worth noting, though, that the discounter's dominance problem runs deeper than the orthodox decision theorist's. While expected utility theory may draw ties between options in cases of *strong* dominance, it still respects the more modest principle of *Strict* Statewise Dominance: if X is *certain* to turn out strictly better than Y, then X is strictly better than Y. But discounting to zero conflicts even with the strict principle, in combination with (Weak) Ex Ante Pareto. Suppose for concreteness that one in a million is a small enough probability to be rounded down to zero. Then consider a million-ticket fair lottery, with a million ticket-holders; the winner gets a cake, and everyone else gets nothing. By assumption, a one in a million chance of cake is no better for each individual than nothing. So by Ex Ante Pareto, this lottery is no better overall than giving nothing to anyone. But giving one person a cake (for sure) is better than giving no one anything, so by Statewise Dominance, the lottery is better than giving nothing to anyone. Contradiction.

small probabilities not discounted to *zero*, but rather to *infinitesimal* values. Then, even for very small probabilities $p < q$, we can say that chance q of a good x is (infinitesimally) better than chance p of x. This style of discounting theory also satisfies Stochastic Compensation. (We take discounting up further in Section 3.2.)

Even Stochastic Compensation builds in more than we really need. Just as we started by considering numerical utilities, and then moved to more abstract qualitative principles about individual good, we can similarly abstract from the precise numerical *probabilities* we have taken for granted thus far. Instead, we can consider a relation of *comparative* probability between events. It is straightforward to restate Stochastic Compensation in terms of such a relation, rather than the probability function P; and it is also not difficult to state sufficient simple structural constraints on this relation. Roughly, it suffices that each event can be subdivided into two equally probable events—intuitively, we can always toss another coin. This is spelled out precisely as Theorem 2 in Appendix A.

All along, these paradoxes have drawn on treating *individuals* and *states* in a broadly analogous way. The analogy becomes especially sharp when the paradox is cast in these general terms. Ex Ante Pareto and Statewise Dominance are each dominance principles, one applied to a prospect person by person, the other state by state. Interpersonal Compensation and Stochastic Compensation also have parallel structure, for these two dimensions. Interpersonal Compensation tells us that we can shift goods from one set of individuals to another set, while improving a state overall. Stochastic Compensation tells us that we can shift goods from one set of states to another set, while improving things for an individual overall.

We have been exploring a clash between Ex Ante Pareto and Statewise Dominance. We began with a paradox that relied on a strong assumption about trade-offs between individuals (Additivity) and a strong assumption about trade-offs involving risk (Individual Expectations). Now we have found that problem is not specific to particular ways of aggregating individual goods or aggregating risks—it is not just a problem for utilitarians, nor for orthodox decision theorists. It is a problem for almost any view that takes *impartial moral good* seriously: that there is such a thing as what is better overall, in the face of risk, and that it is in harmony with what is better for individuals, and with what is better objectively.[12]

[12]This paradox for infinite populations has correlates for *large finite* populations (Good-

3 Where Does This Leave Us?

There are few places left to turn. There remain four kinds of consistent view, classified by which principles they give up:

- Statewise Dominance
- Stochastic Compensation
- Interpersonal Compensation
- Ex Ante Pareto

Table 3 summarizes the views we discuss in this section, and which principles they give up.

3.1 Rejecting Statewise Dominance

Some have defended ethical views that violate *Statewise Dominance* (Arntzenius, 2014; Bostrom, 2011). But this was a bug, not a feature, and we agree with Wilkinson (2022) that it is a pretty devastating problem for those views. As many others have noted, what we really care about when we evaluate uncertain prospects, and rank some as better than others, is achieving better *outcomes*. If we don't know for certain which prospect has a better outcome, this can be tricky. But in cases where we *do* know for certain which prospect has a better outcome, then the question is easy. A theory of betterness-for-prospects that does not give the right ranking in cases where we already know with certainty which will turn out better seems to be just changing the subject. We won't take this up any further.

3.2 Rejecting Stochastic Compensation

Consider a simple gamble X with chance p of a good thing x. Giving up Stochastic Compensation means saying that another simple gamble with

---

sell, 2021). In *Two Teams*, in each state only *finitely* many individuals get a non-zero good. If we interpret the "zero" outcome for each individual as *never having been born*, then each outcome results in a finite population. But Interpersonal Compensation and Ex Ante Pareto must be reinterpreted in this setting, and they have quite different statuses when applied across different populations.

	Statewise Dominance	Stochastic Compensation	Compensation	Ex Ante Pareto
Arntzenius 2014	✗			
Bostrom 2011	✗			
Discounting on a privileged partition		✗		
Social discounting			✗	
Pure Pareto			✗	
Qualitative distribution views				✗
Ex post Paretian views				✗

Table 3: Some examples of axiological theories and the principles they satisfy. Each view satisfies all of the principles except the ones marked with an X.

chance at least p of something at least as good as x is not as good as X. On its face, this seems like an absurd premise for a decision theory.

Even so, there *are* decision theories that have this upshot. These theories tell us to disregard certain possibilities in a way that is not captured just by their probabilities, but rather by other features. One *externalist* idea is that we can ignore states that are sufficiently *distant* from the actual state (as in Bacon, 2014; see Hong, 2024). States where long sequences of coin flips all come up tails may be regarded as especially distant—even while other, equally improbable, sequences of coin flips are closer, and not ruled out. Another related idea would be to rank states based on their *normality* (as in Goodman & Salow, 2023), and disregard states that are sufficiently weird—where this is also understood in a way that cuts across improbability. There are also ways of discounting small probabilities that violate Stochastic Compensation. In general, it makes no sense to disregard *all* propositions with very small probability—such propositions jointly exhaust all of logical space (compare Hájek, 2014, p. 561). So one might introduce some *privileged partition* of states, and disregard cells of that partition with negligibly small probabilities (as in Hong, 2023; Goodman & Salow, 2021).

Here's how this might go in the Two Teams example. In this case, we might naturally regard the number of St. Petersburg Flips that come up tails before the first occurrence of heads as a measure of the distance or weirdness of a state. So, for large n, the states HH_n and TH_n are counted as further or stranger as n increases. For this example, this recipe amounts to the same thing as privileging the partition of states corresponding to the columns in Table 2, and ignoring sufficiently improbable columns. Then according to any of these recipes, the two prospects are to be evaluated in a way that only takes into consideration the states $HH_1, TH_1, HH_2, TH_2, \ldots, HH_n, TH_n$, up to some particular number n.

Evaluating prospects in this way violates Stochastic Compensation. Consider a prospect that provides you with a prize just in the event HH_n. Then consider another prospect that provides you with a prize just in case either HH_{n+1} or TH_{n+1} obtains. This disjunctive event is just as likely as HH_n. But the states where the St. Petersburg sequence take $n+1$ flips to resolve is counted as weirder than when it only takes n flips, and this crosses the threshold of weirdness that tells us not to take these states seriously at all. So any prize you might receive in the event $HH_{n+1} \cup TH_{n+1}$ gets counted as negligible, and so it cannot outweigh the prospect of getting a prize in the equally likely event HH_n.

This is exactly how things go in the example Two Teams. If n is the threshold, then for all the people in any cohort other than the nth, discounting weird states makes no difference to the value of either the prospect One Team Wins or Both Teams Can Win. But many people (2^{n^2} of them) are right on the line: One Team Wins has a *non-negligible* possibility of giving them a cake, while Both Teams Can Win provides only a *negligible* possibility of cake. Thus One Team Wins is counted as strictly better for these people than Both Teams Can Win, and thus there is no longer any conflict between Dominance and Ex Ante Pareto.

But this is a very strange situation. Violating Stochastic Compensation is already strange. Suppose Namaan is one of the individuals in the nth cohort. For either prospect One Team Wins or Both Teams Can Win, what Namaan gets is determined by the outcomes of $n + 2$ coin flips (the Team-Choosing Flip and the first $n + 1$ St. Petersburg Flips). For each of these two prospects, there are exactly two ways these $n + 2$ coin flips can go that give Namaan a cake. Even so, the views we are exploring say that the two prospects are not equally good for Namaan. Some of these $n + 2$ coin flips count for more than others in determining which prospect is better for him. (This is because, by hypothesis, the Team-Choosing Flip does not contribute to a state's weirdness in the same way as the St. Petersburg Flips.)

We can press this strangeness further by considering a variant of Both Teams Can Win—call it *Even Better*. This goes just like Both Teams Can Win if the St. Petersburg sequence of coin flips gets heads at any point up to the nth flip, right at the threshold for non-negligibility. If the St. Petersburg Flips have not come up heads yet at that point, everyone gets a hundred cakes and the game is over. One Team Wins gives Namaan chance $1/2^{n+1}$ of one cake, while Even Better gives him *twice* that probability of a *hundred* cakes. Even so, these views say that One Team Wins is *better* for Namaan than Even Better. While the probability that Even Better pays out is small, the even *smaller* probability event HH_n in which One Team Wins pays out is counted as non-negligible. This verdict is hard to accept.

3.3 Rejecting Interpersonal Compensation

Interpersonal Compensation doesn't say anything about risk: it is a principle about axiology—about which outcomes are better than which with respect to the distribution of goods among individuals. It says, roughly, that any loss to finitely many individuals can be made up for by some gains (perhaps much greater) to finitely many other individuals (perhaps many more of them). As we have discussed, Interpersonal Compensation is a consequence of utilitarianism, but it is a much more broadly held commitment than that. For example, it is also a consequence of natural *prioritarian* views, which give extra weight to the well-being of the worst off. Interpersonal Compensation only requires that it is better to improve the lot of *many more* of the worst off by at least as much, rather than benefitting fewer badly off people by less. Likewise, Interpersonal Compensation follows from natural versions of *sufficientarianism*, which say it is only good to benefit people beyond some threshold if everyone else at least reaches it. We can choose the baseline outcomes for individuals to be above the threshold of sufficiency for everyone. Indeed, giving up Interpersonal Compensation leads to quite radical axiologies.

Utilitarianism, prioritarianism, sufficientarianism, and many other ethical views share in common the idea that no one matters intrinsically more than anyone else. In particular, an overall distribution of goods is just as good if you switch one person's well-being with another. (This idea requires us to make sense of two individuals having the *same* level of well-being—something we have not assumed thus far.) Let's make this principle precise. Say that social outcomes x and y *swap* individuals i and j iff $x(k) = y(k)$ for all individuals k besides i and j, and $x(i) = y(j)$ and $x(j) = y(i)$.

Finite Anonymity. For any social outcomes x and y that swap a pair of individuals, $x \sim y$.

This implies Interpersonal Compensation: if Finite Anonymity holds, then we can always compensate for the loss of some goods to some individuals by giving the very *same* goods to other individuals.[13] So any consistent

[13]Interpersonal Compensation officially requires that for each finite allocation there is some allocation to other individuals which is *strictly better*. We can get this conclusion using a Pareto principle that we have not officially stated so far:

ethical theory that satisfies Statewise Dominance, Ex Ante Pareto, and Stochastic Compensation must be non-anonymous: switching who gets what can matter.[14]

There are two different ways this could go. Suppose that we have an allocation x, and another allocation y that swaps the goods of two individuals Alice and Bea. If the allocation where Alice gets something better than Bea is *better* than the reverse, then it looks like Alice simply matters more than Bea: Alice's loss outweighs Bea's identical gain. The other option is that, while the two allocations are not equally good, neither allocation is better than the other—that is, they are incomparable. Let's briefly consider views of each kind.

There is a very natural anti-egalitarian theory, which, while quite unpopular among philosophers, has many advocates among economists: *social discounting* (for overview, see Greaves, 2017). Suppose that each individual's goods can be represented by numbers in the interval $[0, 1]$; call these numbers *utilities*. In order to calculate the general good of a distribution, rather than adding up all of the individual utilities (which may give infinite results), we multiply each individual i's utility by a *discount factor* a_i, and then add up the discounted utilities. If the discount factors are chosen so that they all add up to a finite value, then the sum of everyone's discounted utilities will also be finite. Call this sum the *social utility* of an outcome. Then we can compare distributions straightforwardly via their social utilities. It is also easy to extend this to a theory of risky prospects: since social utilities are also bounded, each social prospect is guaranteed to have a well-defined *expected* social utility. The resulting comparisons will satisfy Statewise Dominance and Ex Ante Pareto, as well as not just Stochastic Compensation but also Individual Expectations.

As we noted, social discounting is extremely unpopular with philosophers — and for good reason, since it seems wildly unfair. If individual

Strong Ex Post Pareto. For social outcomes x and y, if $x \succsim_i y$ for each individual i and $x \succ_i y$ for some individual i, then $x \succ y$.

But very similar impossibility results hold for a *Weak* Compensation principle about allocations that are *at least as good*, if we use *Weak* Statewise Dominance and *Strict* Ex Ante Pareto, instead of the other way around as we have done.

[14] It might be tempting to generalize Finite Anonymity to arbitrary permutations of individual goods, rather than just swaps. But this strong anonymity principle is inconsistent with Ex Ante Pareto on its own, and indeed with even more modest principles (Hamkins & Montero, 2000).

discount factors are to add up to a finite value, then, first, they cannot be equal (which is why Finite Anonymity fails). Moreover, individuals' discount factors must tend toward zero in the limit. This is how the theory avoids Interpersonal Compensation (and thus the impossibility result): in fact, for any allocation that gives one good (however small!) to a single individual, there is some *infinite* collection of individuals such that giving the best good (with utility 1) to *all* of them is worse. This does seem rather scandalous.

The alternative to inequity is incomparability: perhaps swapping the goods of two individuals does not make things better or worse, but also does not leave things equally good. The simplest axiology like this is the *pure Paretian* view, which says that one distribution is as good as another if *and only if* it is as good for every individual. This takes anti-Compensation to its extreme: no loss to anyone can be compensated by any gains to anyone else. One natural motivation for this view is the idea that there is no way of comparing goods for different individuals, and so no way of judging whether a gain to one individual is greater or smaller than a corresponding loss to another. (This motivation implies that Finite Anonymity is not even wrong, but rather nonsense.)

It is again straightforward to extend the pure Paretian view to a theory of risk. If we suppose again that each individual's good can be represented by utilities between 0 and 1, then we can consider the *pure ex ante Pareto* order on prospects, which says that one prospect is better than another if and only if it has higher expected utility for every individual. This also satisfies Statewise Dominance.

The main cost of the pure Paretian view is that betterness gives us almost no guidance whatsoever. Almost anything non-trivial we might do will have at least very small negative effects on someone. For example, this view says that saving a drowning child at the trifling personal cost of muddying one's shoes does not make the world better. It also faces general objections to incomparability with respect to betterness, which we will not rehearse.[15]

[15]For example, it is inconsistent with the combination of Stochastic Equivalence (the principle that prospects with the same probability distribution over outcomes are equally good) with the following (see Schoenfield, 2014):

Negative Dominance. For any prospects X and Y, if X is no better than Y in every state, then X is not better than Y.

There isn't much room for anti-Compensation views that are very different in spirit from both social discounting and pure Paretianism. Either different individuals are treated differently from one another, or else it is quite difficult to compensate for any loss to some people by any gain to others. Let's make this more precise.

Even though pure Paretianism does not uphold Finite Anonymity, it still treats different individuals exactly the same: its betterness ordering is *invariant under permutations of individuals*. If we uniformly rearrange who gets what, this makes no difference to which social outcomes are better (for discussion see Askell, 2018). In contrast, social discounting does not have this symmetry. Moreover, any permutation-invariant axiology that violates Interpersonal Compensation must be quite similar to pure Paretianism. Call a social outcome that just gives each individual either a cake or nothing *simple*. If Interpersonal Compensation fails for such outcomes, then this holds:[16]

Almost Pure Pareto. There is some number n such that, for any simple outcomes x and y, if x is worse than y for at least n individuals, then x is not as good as y overall.

The pure Paretian view corresponds to the case where $n = 1$. In principle, Almost Pure Pareto could fail for $n = 1$ while holding for some larger number: perhaps a loss to one individual can be outweighed by gains to others, but no loss to *six* individuals can be outweighed by gains to others, however many. But it is hard to imagine a motivation for this kind of view.

In short, any view that gives up Interpersonal Compensation will have the same general flavor as either social discounting—which treats individuals unequally—or pure Paretianism—which allows very few betterness comparisons.

3.4 Rejecting Ex Ante Pareto

Pareto principles in ethics have faced a wide variety of challenges before this one. It is worth recalling at the outset, though, that the principle we are discussing is not a principle about *preference*, but about *betterness.* Because of this, some standard objections do not straightforwardly apply.

For other arguments, see Dorr et al. (Manuscript).

[16] This also relies on Strong (Ex Post) Pareto.

For instance, it is well-known that there is no social preference relation that always respects the unanimous preference of individual expected utility maximizers who not only have different utilities, but also have conflicting *credences* (for discussion see Broome, 1991, p. 152). But what is best for an individual with respect to certain probabilities can come apart from what that individual prefers, if they do not share those probabilities. So long as we are careful to state our Ex Ante Pareto principle about what is better for individuals with respect to a single fixed probability function (as we have been), this kind of objection has no bite. Similar things might be said about Blessenohl (2020)'s argument concerning individuals with different attitudes toward risk, and perhaps also the objection from Mongin (2016) concerning individuals whose preferences are based on different private reasons that a social planner is not bound to respect. Even after taking on board the morals of these other challenges, it still seems really odd to say that a prospect might be worse overall while being *better* for every individual.

Still, some philosophical views do have this upshot. Some reject the moral importance of *people* altogether. Note that the Ex Ante Pareto principle relies on some privileged way of identifying the same *individual person* in different possible states of nature—but in general this may be hard to do, and it may be hard to see such an identification as morally significant (compare Parfit, 1984, p. 215; for a different, somewhat related difficulty, see Mahtani, 2017). Perhaps what really matters is the distribution of good and bad *experiences*, or satisfied *preferences*, or *states of flourishing*. Individual people, on this kind of view, are mere receptacles of value without special significance of their own.

This kind of reason for rejecting Ex Ante Pareto harmonizes with also rejecting the *Ex Post* Pareto principle, which says that an *outcome* which is better for each individual is better overall. Take a simple example from Hamkins and Montero (2000): suppose individual welfares can be represented by integers, and there is one person with each integer welfare. Now consider another outcome in which every person is better off by one. Both outcomes have exactly the same qualitative distribution of welfares. If individual people are not central to moral value, but rather general distributions, this suggests that the two outcomes are equally good, even though one is better for every individual person. Rejecting Pareto principles provides a unified response to a variety of such puzzles (such as Cain, 1995; Askell, 2018).

This is not the only option, though. There are consistent views that give up Ex Ante Pareto, but keep Ex Post Pareto.[17] Roughly speaking, this kind of view says that how things turn out for individual people matters to what is best overall, but *risks* to individual people do not. Indeed, the same kind of argument as in Two Teams shows that a prospect that imposes some risk of harm on every individual can be strictly worse than a prospect that imposes a *larger* risk of the same harm on every individual. Then we face the unsettling question of when reducing risks of harm to individuals is good, and when it really makes things worse.

4 Conclusion

We began this paper with a paradox that threatened a combination of quite specific views: a utilitarian axiology that weighs individual welfare by simply adding it up—where individual utility functions were assumed to be unbounded—and an expectational decision theory that weighs prospects by a weighted average of outcomes. We stripped away these assumptions, revealing a paradox at the core which threatens a much wider range of ethical views. Putting it a bit roughly, this includes nearly *any* view that has a place for comparisons of overall moral value of social prospects.

Any adequate response to the paradoxes is going to cut deep. To sum up the options briefly, and a bit roughly:

- We might give up Statewise Dominance. Then we must give up the idea that "subjective" value (that is, overall betterness for prospects) is a guide to "objective" value (that is, overall betterness for outcomes). For sometimes we might know with certainty that one prospect is objectively better, but still regard it as subjectively worse.

- We might give up Stochastic Compensation. This would be to say that a *larger* chance at some benefit may not be as good for you as a *smaller* chance of the same benefit.

- We might give up Interpersonal Compensation. This will involve either giving up on the idea that there can be moral trade-offs between

[17]For discussion see Wilkinson (2022); note, though, that Wilkinson does not prove that the principles he offers are consistent.

persons at all—pushing us toward the pure Paretian view—or else giving up on the idea that different people's interests ought to be given equal moral weight.

- We might give up Ex Ante Pareto, giving up on the idea that betterness for individual people plays a central role in overall betterness, at least when it comes to risk.

Whichever way we go, serious revision of our ordinary moral views is called for. We do not regard these paradoxes of infinity as mere theoretical puzzles. It is, in fact, a serious possibility that we live in the kind of infinite world on which they bear directly.

A Proofs of Theorems

First a few technical preliminaries. We consider a set $\mathcal{S}$ of *states* equipped with a σ-algebra $\mathcal{E}$ of *events*, and a probability measure P on this algebra. We assume that P is suitably rich: in particular, that it is non-atomic. (For any event E and any probability $0 < p < P(E)$, there is some event $F \subseteq E$ such that $P(F) = p$.)

There is a countably infinite set $\mathcal{I}$ if *individuals*. For each $i \in \mathcal{I}$ there is a set $\mathcal{O}_i$ of *individual outcomes* for i (or *goods*). There is some "baseline" outcome $0_i \in \mathcal{O}_i$.

A *social outcome* is a function that takes each individual i to an individual outcome for i. Recall that a social outcome x is an *allocation for* a set of individuals I iff $x(j) = 0_j$ for each individual $j \notin I$. (An allocation for I is also an allocation for any $J \supset I$.) If x and y are allocations for disjoint sets of individuals I and J, respectively, let $x \sqcup y$ be the piecewise-defined allocation for $I \cup J$,

$$(x \sqcup y)(i) = \begin{cases} x(i) & \text{for } i \in I \\ y(i) & \text{for } i \in J \\ 0_i & \text{otherwise} \end{cases}$$

We can similarly define an infinitary operation $x_1 \sqcup x_2 \sqcup \cdots$.

A *social prospect* is a (measurable) function from states to social outcomes.[18] There is a relation $\succsim$ on social prospects ("at least as good"). We

[18]Officially, $\mathcal{O}_i$ should also be equipped with a σ-algebra, and the set of social outcomes

assume this is transitive and reflexive (a *preorder*). The relations $\prec$ and $\sim$ are defined in terms of $\precsim$ in the standard way.

An *individual prospect* for $i \in \mathscr{I}$ is a (measurable) function from states to O_i. There are preorders $\succsim_i$ on individual prospects for each individual i. We also write $X \precsim_i Y$ for an individual i and *social* prospects X and Y, to indicate that $X_i \precsim_i Y_i$, where X_i is the individual prospect such that

$$X_i(s) = X(s)_i \quad \text{for each } s \in \mathscr{S}$$

We will often let a social/individual outcome x stand in for the constant prospect that achieves x in every state, when there is no risk of confusion.

Theorem 1. *(Weak) Ex Ante Pareto, (Strict) Statewise Dominance, Interpersonal Compensation, and Stochastic Compensation are jointly inconsistent.*

Proof. Let $E_1, E_2, \ldots, F_1, F_2, \ldots$ be an infinite partition of non-null events with the property that for each n,

$$P(E_n) = P(F_n) \leq P(E_{n+1} \cup F_{n+1})$$

Concretely, we can let $P(E_n) = P(F_n) = 1/2^{n+1}$. We will construct prospects X and Y with the structure shown in Table 4.

First, we recursively define four sequences of allocations $w_1, w_2, \ldots$, $x_1, x_2, \ldots, y_1, y_2, \ldots$, and $z_1, z_2, \ldots$.

For the base case, let $x_0 = w_0 = 0$.

For the recursive step, let $k > 0$ and let $I_1, J_1, \ldots, I_{k-1}, J_{k-1}$, x_{k-1}, and w_{k-1} be given. By Compensation, there is some allocation y_k for a finite set of individuals I_k disjoint from $I_1 \cup J_1 \cup \cdots \cup I_{k-1} \cup J_{k-1}$ such that

$$x_{k-1} \sqcup w_{k-1} \prec y_k$$

Likewise, there is an allocation z_k for a finite set of individuals J_k disjoint from $I_1 \cup J_1 \cup \cdots \cup I_{k-1} \cup J_{k-1} \cup I_k$ such that

$$x_{k-1} \sqcup w_{k-1} \prec z_k$$

By Stochastic Compensation, since $P(E_k) \leq P(E_{k+1} \cup F_{k+1})$, there is also an allocation x_k for I_k such that

$$y_k|_{E_k} \prec_i x_k|_{E_{k+1} \cup F_{k+1}}$$

is endowed with the product algebra. But this is unimportant, since our focus will be on discrete prospects.

X	E_1	F_1	E_2	F_2	E_3	F_3	$\cdots$
I_1			x_1	x_1			
J_1			w_1	w_1			
I_2					x_2	x_2	
J_2					w_2	w_2	
$\vdots$							$\ddots$

Y	E_1	F_1	E_2	F_2	E_3	F_3	$\cdots$
I_1	y_1						
J_1		z_1					
I_2			y_2				
J_2				z_2			
I_3					y_3		
J_3						z_3	
$\vdots$							$\ddots$

Table 4: The proof of Theorem 1

Likewise, there is an allocation w_k for J_k such that

$$z_k|_{E_k} \prec_i w_k|_{E_{k+1} \cup F_{k+1}}$$

Now we construct prospects X and Y such that X has outcome $x_k \sqcup w_k$ in each event $E_{k+1} \cup F_{k+1}$, while Y has outcome y_k in each event E_k and z_k in each event F_k. We can then reason as before: by construction, $X \prec Y$ by Statewise Dominance, while $Y \precsim X$ by Ex Ante Pareto. □

Finally, it is straightforward to reformulate things in terms of a qualitative relation of comparative probability between events, which we will write $\sqsubseteq$. We will assume without further comment that $\emptyset \sqsubseteq \mathcal{S}$. But we rely on no additional assumptions—not even that $\sqsubseteq$ is transitive.

Comparative Stochastic Compensation. For any good x for an individual i, and any events E and F such that $E \sqsubseteq F$, there is some good y for i such that $x|_E \precsim y|_F$.

It is clear from inspection that the following structural condition on $\sqsubseteq$ suffices for the proof of Theorem 1.

Partition There exists an infinite partition of events $E_1, E_2, \dots, F_1, F_2, \dots$ such that, for each n,

$$E_n \sqsubseteq E_{n+1} \cup F_{n+1}$$
$$F_n \sqsubseteq E_{n+1} \cup F_{n+1}$$

For this, the following suffices:

Decomposition If E and F are disjoint events such that $E \sqsubseteq F$, then there are events E_1 and E_2 that partition E and events F_1 and F_2 that partition F, where

$$E_1, E_2 \sqsubseteq F_1 \sqsubseteq F_2$$

Intuitively, E_1 (E_2) might correspond to "E and a coin toss comes up heads (tails)", and similarly for F_1, F_2 and F—where the coin toss is understood to be independent of E and F.

We can then recursively construct the required sequence for Partition. First, apply Decomposition to $\emptyset$ and $\mathcal{S}$ to get events $A_0 \sqsubseteq B_0$ that partition $\mathcal{S}$. Then, for each n, given events $A_n \sqsubseteq B_n$ we can apply Decomposition to get events E_n and F_n that partition A_{n-1} and events A_n and B_n that partition B_{n-1}, where

$$E_n, F_n \sqsubseteq A_n \sqsubseteq B_n$$

By construction $A_n = E_{n+1} \cup F_{n+1}$, so we have the required properties for Partition.

Theorem 2. *Ex Ante Pareto, Statewise Dominance, Interpersonal Compensation, Comparative Stochastic Compensation, and Decomposition are jointly inconsistent.*

Acknowledgements

Early drafts of this paper benefitted greatly from discussions with Andrew Bacon, Zach Goodsell, John Hawthorne, Jake Nebel, and Jake Ross. Thanks also to several anonymous referees for their feedback.

References

Arntzenius, F. (2014). Utilitarianism, decision theory and eternity. *Philosophical Perspectives*, *28*(1), 31–58. https://doi.org/10.1111/phpe.12036

Askell, A. (2018). *Pareto principles in infinite ethics* (Doctoral dissertation). New York University.

Bacon, A. (2014). Giving your knowledge half a chance. *Philosophical Studies*, (2), 1–25. https://doi.org/10.1007/s11098-013-0276-6

Basu, K., & Mitra, T. (2007). Utilitarianism for infinite utility streams: A new welfare criterion and its axiomatic characterization. *Journal of Economic Theory*, *133*(1), 250–373.

Beckstead, N., & Thomas, T. (2021). A paradox for tiny probabilities and enormous values. *Global Priorities Institute Working Paper*, (7-2021).

Blessenohl, S. (2020). Risk attitudes and social choice. *Ethics*, *130*(4), 485–513. https://doi.org/10.1086/708011

Bostrom, N. (2011). Infinite ethics. *Analysis and Metaphysics*, *10*, 9–59.

Broome, J. (1991). *Weighing goods: Equality, uncertainty and time*. Wiley-Blackwell.

Buchak, L. (2013). *Risk and rationality*. Oxford University Press.

Cain, J. (1995). Infinite utility. *Australasian Journal of Philosophy*, *73*(3), 401–404. https://doi.org/10.1080/00048409512346731

Dorr, C., Nebel, J. M., & Zuehl, J. (Manuscript). *The case for comparability*. https://philpapers.org/rec/DORTCF-2

Easwaran, K., Hájek, A., Mancosu, P., & Oppy, G. (2021). Infinity. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2021). Metaphysics Research Lab, Stanford University.

Goodman, J., & Salow, B. (2021). Knowledge from probability. In J. Y. Halpern & A. Perea (Eds.), *Theoretical aspects of rationality and knowledge tark (2021)*. Oxford University Press.

Goodman, J., & Salow, B. (2023). Epistemology normalized. *Philosophical Review*, *132*(1), 89–145. https://doi.org/10.1215/00318108-10123787

Goodsell, Z. (2021). A st petersburg paradox for risky welfare aggregation. *Analysis*, *81*(3), 420–426. https://doi.org/10.1093/analys/anaa079

Greaves, H. (2017). Discounting for public policy: A survey. *Economics and Philosophy*, *33*(3), 391–439. https://doi.org/10.1017/s0266267117000062

Hájek, A. (2014). Unexpected expectations. *Mind*, *123*(490), 533–567. https://doi.org/10.1093/mind/fzu076

Hamkins, J. D., & Montero, B. (2000). With infinite utility, more needn't be better. *Australasian Journal of Philosophy*, *78*(2), 231–240. https://doi.org/10.1080/00048400012349511

Hong, F. (2023). Prefaces, knowledge, and questions. *Ergo: An Open Access Journal of Philosophy*, *10*(n/a). https://doi.org/10.3998/ergo.4631

Hong, F. (2024). Know your way out of st. petersburg: An exploration of knowledge-first decision theory. *Erkenntnis*, 1–20. https://doi.org/10.1007/s10670-022-00639-2

Horton, J. (2017). Aggregation, complaints, and risk. *Philosophy and Public Affairs*, *45*(1), 54–81. https://doi.org/10.1111/papa.12084

Isaacs, Y. (2016). Probabilities cannot be rationally neglected. *Mind*, *125*(499), 759–762. https://doi.org/10.1093/mind/fzv151

Kowalczyk, K. (n.d.). *Risk times infinity*.

Lauwers, L., & Vallentyne, P. (2004). Infinite utilitarianism: More is always better. *Economics and Philosophy*, *20*(2), 307–330. https://doi.org/10.1017/s0266267104000227

Mahtani, A. (2017). The ex ante pareto principle. *Journal of Philosophy*, *114*(6), 303–323. https://doi.org/10.5840/jphil2017114622

Mongin, P. (2016). Spurious unanimity and the pareto principle. *Economics and Philosophy*, *32*(3), 511–532. https://doi.org/10.1017/S0266267115000371

Monton, B. (2019). How to avoid maximizing expected utility. *Philosophers' Imprint*, *19*.

Parfit, D. (1984). *Reasons and persons*. Oxford University Press.

Parfit, D. (1986). Overpopulation and the quality of life. In P. Singer (Ed.), *Applied ethics* (pp. 145–164). Oxford: Oxford University Press.

Scanlon, T. (1998). *What we owe to each other*. Belknap Press of Harvard University Press.

Schoenfield, M. (2014). Decision making in the face of parity. *Philosophical Perspectives*, *28*(1), 263–277. https://doi.org/10.1111/phpe.12044

Smith, N. J. J. (2014). Is evaluative compositionality a requirement of rationality? *Mind*, *123*(490), 457–502. https://doi.org/10.1093/mind/fzu072

Tarsney, C. (2020). Exceeding expectations: Stochastic dominance as a general decision theory. *Global Priorities Institute Working Paper*, (3).

Vallentyne, P., & Kagan, S. (1997). Infinite value and finitely additive value theory. *Journal of Philosophy*, *94*(1), 5–26. https://doi.org/jphil199794115

Wilkinson, H. (2022). Infinite aggregation and risk. *Australasian Journal of Philosophy*. https://philpapers.org/rec/WILIAA-17